



Comparison between image texture and polarization features in histopathology

YUDI LIU,^{1,5} YANG DONG,^{1,5} LU SI,¹ RUOYU MENG,² YANMIN DONG,³ AND HUI MA^{1,2,4,*}

¹Center for Precision Medicine and Healthcare, Tsinghua-Berkeley Shenzhen Institute, Shenzhen 518055, China

²Shenzhen Key Laboratory for Minimal Invasive Medical Technologies, Institute of Optical Imaging and Sensing, Graduate School at Shenzhen Tsinghua University, Shenzhen 518055, China

³Shenzhen Hospital of Traditional Chinese Medicine, Shenzhen 518034, China

⁴Department of Physics, Tsinghua University, Beijing 100084, China

⁵These authors contributed equally to this work

*mahui@tsinghua.edu.cn

Abstract: Digital pathology has shown great importance for diagnostic purposes in the digital age by integrating basic image features into multi-modality information. We quantify the degree of correlation between the multiple texture features from H&E images and polarization parameter sets derived from Mueller matrix images of the same sample to provide more microstructural information for assisting diagnosis. The experimental result shows the correlations between texture feature and polarization parameter via Pearson coefficients. Polarization parameters t_1 , D_L and the depolarization parameter Δ correlated with image texture features *Tamura_Fcon* and *Tamura_Frgh*, and can be used as powerful tools to quantitatively characterize cell nuclei related with tumor progression in breast pathological tissues. Polarization parameters δ and r_L associated with the image texture feature *Tamura_Flin* have great potential for the quantitative characterization of proliferative fibers produced by inflammation. Furthermore, polarization parameters have the advantages of stable recognition in low resolution images. This work validates the associations between image texture features and polarization parameters and the merit of polarization imaging methods in low-resolution situations.

© 2021 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Digital pathology is the endorsement for pathologists in the diagnosis of breast cancers. Digital image processing methods are able to enhance the image quality for better quantifying repeated patterns especially for texture analysis. Quantitative characterization that can provide repeatable, objective and reliable results are essential. Pathological analysis of breast tumors and assessment of prognosis is based on accurate segmentation and quantitative evaluation of cell nuclei [1] and alignment of collagen fiber [2] in histopathological images. For examples, Lu *et al.* [3] proposed a quantitative measure of cellular texture characteristics and investigated for contributing to automatic classification between different cell nuclei types in breast cancer. Jones *et al.* [4] quantitatively characterized abundance and alignment of intra-tumoral collagen from polarimetric images and used linear discriminant model to predict prognostic on invasive ductal carcinoma. And Chen *et al.* [5] quantified collagen morphology in the second harmonic generation (SHG) image via Tamura texture features in local orientation ternary pattern. The gray-level run-length matrix-based features, co-occurrence matrix based-features and Tamura texture features contributed most in this study.

Mueller matrices contain a wealth of microstructural information and optical properties when detecting complex biological samples [6,7]. There are several methods for mining information from Mueller matrices such as Mueller matrix polar decomposition (MMPD) [8], Mueller

matrix transform (MMT) [9] and logarithmic Mueller matrix decomposition (LMMD) [10,11]. Polarization parameters with clear physical meanings were derived from Mueller matrices have shown good diagnostic potential in characterizing the microstructure of various cancer pathological tissues, including skin cancer [12], breast cancer [13–15], cervical cancer and thyroid carcinoma [16]. In addition, in previous studies, based on machine learning methods, we used the Mueller matrix parameters to construct polarimetry feature parameters that can specifically quantify multiple microstructures in different breast tissues [13].

At high resolution, the correspondence between texture features and polarization parameters can provide preliminary prediction of the polarimetry characterization results when obtaining H&E-stained microscopic images without measuring the Mueller matrix images using special polarization imaging microscopy. In the previous studies [13], the target microstructures could be specifically and quantitatively identified by the polarimetry feature parameters (PFPs) composed of polarimetry basis parameters (PBPs). Therefore, studying the correspondence between texture features and polarization parameters can provide a theoretical basis for direct use of image texture features to guide the primary screening of PBPs in later studies. In previous studies, our group has demonstrated that the polarization imaging method is less resolution-dependent in quantitatively characterizing wavelength-scale microstructures [13,17]. Therefore, once the correspondence between image texture features and polarization parameters is established, the polarization parameters can provide information with clear physical meaning that is missing from H&E images supplementally. In the case of insufficient resolution resulting in the failure of the texture features to identify the target microstructure.

In this paper, microscopic Mueller matrix images of H&E pathology sections of breast ductal carcinoma were obtained using Mueller matrix microscopy. Images of polarization parameters from MMPD and MMT [9,18,19] were then calculated. RGB images of H&E pathology sections were analyzed using a 50*50 pixel-size sliding window for multiple texture features to obtain texture feature images. The texture feature image is threshold-segmented, and the selected pixel locations are mapped onto both the texture feature image and the polarization parameter image to obtain an array of pairs corresponding to the pixel locations, and quantitative correlation analysis is performed. In the following sections, we describe our method, show results for characterizing different target microstructures using texture features and polarization parameters, and demonstrate texture Feature-polarization parameter correlations via heat maps [20]. Image texture features *Tamura_Fcon* and *Tamura_Frgh* correlated with depolarization parameter Δ , polarization parameters of t_1 and D_L have great potential for the characterization of cell nuclei in human breast ductal tissue. Image texture feature *Tamura_Flin* associated with polarization parameters of δ and r_L to quantitatively identify fiber tissues in breast ductal carcinoma pathological slides. In the validation section, to identify target structures, we used H&E pathology slice samples from 15 patients with breast ductal carcinoma to investigate the ability of several one-to-many texture features-polarization parameters, which is pairwise and highly correlated in multi-resolution situations. This experiment highlights the superiority of polarization imaging methods in low resolution, large field of view situations.

2. Method and materials

2.1. Experimental setup and breast duct tissues sample

As shown in Fig. 1, we used a commercial transmission microscope (L2050, Liss Optical Instrument Factory, Guangzhou, China) with subjoining polarization states generator (PSG) module and polarization states analyzer (PSA) module [21]. During the measurement, forward scattered light contained polarization information is captured by grayscale CCD, and RGB image of pathological slide is collected by color CCD. Polarization intensity images are collected every 6° of rotation of R2, and every 30° of rotation of R1. In each measurement, R2 rotated 180°, therefore 30 intensity images with different polarization states were captured. The 5:1 rotation

ratio of $R1(5\omega)$: $R2(\omega)$ is the lowest ratio in which the expressions for the Fourier coefficients can be inverted to give the Mueller matrix elements. 16 Mueller matrix elements of samples can be calculated by the typical dual-rotating retarder method [22]. After collecting the Mueller matrices of the air, the system is calibrated by the numerical calibration method [23], and the maximum error is less than 0.01.

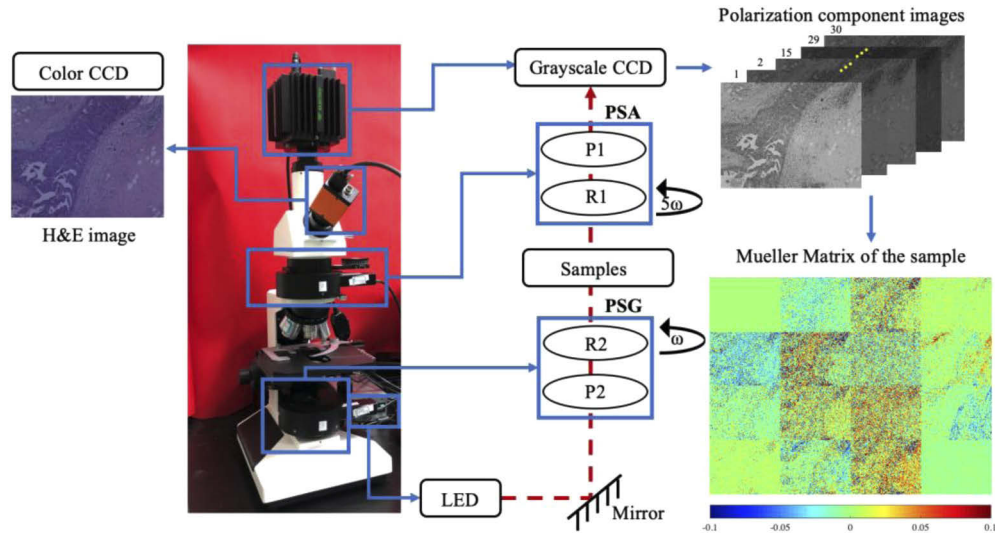


Fig. 1. Schematic of Mueller matrix microscope: LED (Cree, 3W, 633 nm, $\Delta\lambda = 20$ nm), Grayscale CCD (QImaging, 12-bit, 74-0107A, Canada), Color CCD (8-bit, JHSM300f, China), Polarization State Analyzer (PSA): polarizer P1 (extinction ratio 500:1, Daheng Optics, China) and a rotatable quarter-wave plate R1 (Daheng Optics, China), Polarization State Generator (PSG): fixed polarizer P2 (extinction ratio 500:1, Daheng Optics, China) and a rotatable quarter-wave plate R2 (Daheng Optics, China). An example of thirty intensity images and the normalized Mueller matrices (plot with diagonal elements minus 1) of the H&E pathological slide is presented.

In this study, the samples are H&E-stained pathological sections of human breast ductal carcinoma tissues provided by Shenzhen Hospital of Traditional Chinese Medicine. Considering about individual variability between patients, a total of 15 H&E pathological slices was selected from patients diagnosed as breast ductal carcinoma. The 4- μm -thick pathological slices can be considered as a single layer of structure whose integrally averaged polarization properties are all contributed by the target microstructures. Region of Interest (ROI) in each pathological section was selected from each case as sample, thus 15 samples were analyzed in this study. The ROI contains a large number of cancer cells caused by canceration and abundant fiber structures caused by inflammatory reaction, which have great meaning for diagnosis and prognosis. The target microstructures-cell nuclei and fiber tissues-in the breast pathological images are all labelled by experienced pathologists, which can be used as ground truth for the identification results by texture and polarization parameters. This work was approved by the Ethics Committee of Shenzhen Hospital of Traditional Chinese Medicine.

2.2. Polarimetry basis parameters (PBPs)

Previous studies on various pathological tissues have shown that the Mueller matrix imaging method has good results in detecting the wavelength scale microstructures and improving the diagnosis of cancer pathology samples [13,16,24].

Among Mueller matrix decomposition methods, MMPD proposed by Lu and Chipman is adaptable for analysis of transmission Mueller matrix of complex biological tissues [25]. MMT from our previous work have been employed in various pathological samples, and demonstrated its preliminary biomedical application prospects. Therefore, as shown in Table 1, eleven polarimetry basis parameters, including MMPD and MMT parameters, with clear physics meanings derived from Mueller matrices were used in this study to analyze the transmission Mueller matrix of breast pathological samples. MMPD parameters include depolarization (Δ), linear retardance (δ), diattenuation (D), and orientation of fast axis (θ). In MMT, we also proposed various rotation invariant polarization parameters calculated from Mueller matrix elements and related to linear diattenuation (P_L), linear dichroism (D_L), linear retardance (r_L), linear birefringence (q_L), anisotropy degree (t_1), polarizance (b), normalized anisotropy (A) [6,12].

Table 1. Polarimetry basis parameters used in the study^a

| Mueller matrix polar decomposition | Definition | Mueller matrix polar decomposition | Definition |
|------------------------------------|--|------------------------------------|--|
| Δ | $\Delta = 1 - \frac{ tr(M_{\Delta}^{-1}) }{3}$ | D | $D = \sqrt{(m12)^2 + (m13)^2 + (m14)^2}$ |
| θ | $\theta = 0.5 \tan^{-1} \left(\frac{r_2}{r1} \right)$ | δ | $\delta = \cos^{-1} \left\{ \left[\frac{(M_R(2,2) + M_R(3,3))^2}{+ (M_R(3,2) + M_R(2,3))^2} \right]^{\frac{1}{2}} - 1 \right\}$ |
| Mueller matrix transformation | Definition | Mueller matrix transformation | Definition |
| P_L | $P_L = \sqrt{(m21)^2 + (m31)^2}$ | t_1 | $t_1 = \frac{\sqrt{(m22-m33)^2 + (m23+m32)^2}}{2}$ |
| D_L | $D_L = \sqrt{(m12)^2 + (m13)^2}$ | | |
| q_L | $q_L = \sqrt{(m42)^2 + (m43)^2}$ | b | $b = \frac{m22+m33}{2}$ |
| r_L | $r_L = \sqrt{(m24)^2 + (m34)^2}$ | A | $A = \frac{2b \times t_1}{b^2 + t_1^2}$ |

^aIn the Table above, m11-m44 denoted as 16 elements of the Mueller matrix, M_R is the 4*4 retardance sub-matrix and M_{Δ} is the 3*3 depolarization sub-matrix resulting from the Lu-Chipman polar decomposition, tr represents the trace of the matrix, $r1$ and $r2$ are the components of retardance.

2.3. Image analysis methods

Texture features quantitatively characterize the changes in spatial distribution between adjacent pixels. Multiple texture features were extracted from RGB image of pathological slides of breast ductal carcinoma. The extracted features differentiate micro-structures for breast duct tissue which play a significant role in histopathological digitalization and computer-aided diagnosis. Here, we constructed a texture feature set including Tamura features, gray level co-occurrence matrix (GLCM) features, local binary pattern (LBP) and gray level run length matrix (GLRLM) features, which have been widely used in digital pathology [26,27].

2.3.1. Data preprocessing

This research is based on mapping the selected pixels to the polarization parameter image after thresholding the texture feature image to realize the correlation analysis between the texture features and the polarization parameters at the same pixel position. Therefore, the polarization parameter image and texture feature image before threshold and segmentation are required to be registered pixel by pixel. In the study, we used an affine-transformation-based image registration method [28]. After optimizing the iteration step size, initial conditions and the maximum number of iterations, the pixel-level registration of the two images was achieved.

2.3.2. Image texture analyze methods

Tamura Feature. Tamura proposed six texture features based on human vision and intuition [29]. The six texture features are coarseness, contrast, roughness, directionality, line-likeness and regularity, of which the definitions are shown in Table 2.

Table 2. Texture parameters used in the study ^a

| Tamura Features | Definition | GLRLM features | Definition |
|-----------------|--|----------------|--|
| F_{crs} | $F_{crs} = \frac{1}{m \times n} \sum_{i=1}^m \sum_{j=1}^n S_{best}(i, j)$ | SRE | $SRE = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j)/j^2]}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| F_{con} | $F_{con} = \frac{\sigma}{(\alpha_4)^{1/4}}$ | LRE | $LRE = \frac{\sum_{i=1}^m \sum_{j=1}^n j^2 P_d^\theta(i, j)}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| F_{dir} | $F_{dir} = 1 - \frac{1}{m} \sum_p \sum_{\phi \in \omega_p} (\phi - \phi_p)^2 \dot{H}_D(\phi)$ | GLN | $GLN = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j)]^2}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| F_{lin} | $F_{lin} = \frac{\sum_{i=1}^m \sum_{j=1}^n P_{D,d}(i, j) \cos[(i-j)(2\pi/n)]}{\sum_i}$ | RLN | $RLN = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j)]^2}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| F_{reg} | $F_{reg} = 1 - r(\sigma_{crs} + \sigma_{con} + \sigma_{dir} + \sigma_{lin})$ | RP | $RP = \sum_{i=1}^m \sum_{j=1}^n \frac{P_d^\theta(i, j)}{m \times n}$ |
| F_{rgh} | $F_{rgh} = F_{crs} + F_{con}$ | $LGLRE$ | $LGLRE = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j)/i^2]}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| GLCM features | Definition | $HGLRE$ | $HGLRE = \frac{\sum_{i=1}^m \sum_{j=1}^n i^2 P_d^\theta(i, j)}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| Con | $Contrast(d, \theta) = \sum_{i=0}^m \sum_{j=0}^n i - j ^2 P_d^\theta(i, j)$ | $SRLGLE$ | $SRLGLE = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j)/i^2 j^2]}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| Egy | $Energy(d, \theta) = \sum_{i=0}^m \sum_{j=0}^n [P_d^\theta(i, j)]^2$ | $SRHGLE$ | $SRHGLE = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j) i^2 j^2]}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| Cor | $Correlation(d, \theta) = \sum_{i=0}^m \sum_{j=0}^n \frac{ij P_d^\theta(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$ | $LRLGLE$ | $LRLGLE = \frac{\sum_{i=1}^m \sum_{j=1}^n [P_d^\theta(i, j) i^2 j^2]}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |
| Hmg | $Homogeneity(d, \theta) = \sum_{i=0}^m \sum_{j=0}^n \frac{1}{1 + (i-j)^2} P_d^\theta(i, j)$ | $LRHGLE$ | $LRHGLE = \frac{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j) i^2 j^2}{\sum_{i=1}^m \sum_{j=1}^n P_d^\theta(i, j)}$ |

^aHere, S_{best} means the neighborhood size which generates the highest similarity of intensity, σ represents standard deviation and μ is mean value, α_4 is the kurtosis of the intensity histogram, H_D denoted as direction histogram, based on H_D , n_p is number of peaks, ϕ_p is the position of the p-th peak, ω_p is range of the p-th peak and ϕ is direction angle.

Gray-level co-occurrence matrix (GLCM). GLCM characterizes the texture of an image by calculating the spatial relationship of pixel values, and then four statistical features-contrast, energy, correlation and homogeneity-can be derived from this matrix [30].

Local binary patterns (LBP). Figure 2 shows the four LBP operators used in this study. The original LBP operator is defined in the neighborhood of pixels 3*3, with the center pixel of the neighborhood as the threshold [31]. Ojala *et al.* [32] proposed a uniform pattern to reduce the dimensionality of LBP operators. The Uniform LBP solves the problem of too many binary patterns and improves statistics. The Circular LBP [33] adapts to texture features of different scales by improving LBP operator from square neighborhood to circular neighborhood. Rotation LBP is based on the circular LBP. Rotate circular LBP form different angles, then take the center pixel value as the smallest circular LBP feature value.

Gray-level run length matrix (GLRLM). The run-length matrix $Mr(i, j)$ is defined as the number of runs of variables having gray level i and run length j at $0^\circ, 45^\circ, 90^\circ$ and 135° four directions of the image. As shown in Table 2, eleven GLRLM texture features of run-length statistics derived by Galloway [34], Chu *et al.* [35], Dasarthy and Holder [36] were analyzed in this study.

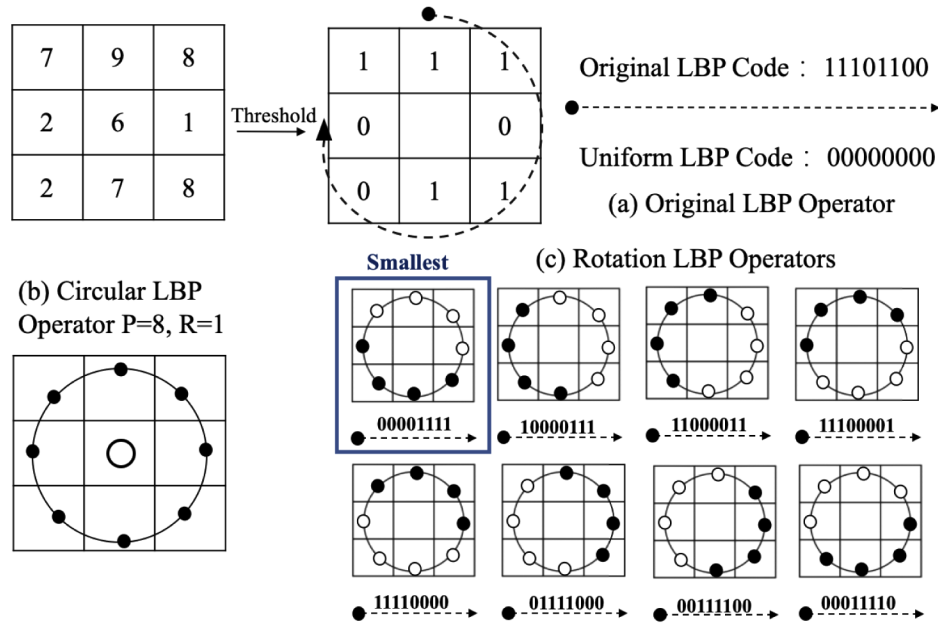


Fig. 2. Summary of LBP Operators. (a) Illustration of original LBP operator. Example of original LBP binary codes obtained by local thresholding and transformed into uniform LBP code. (b) A circular LBP operator with a sampling number P of 8 and a radius of 1 is used in this study. (c) Examples of rotation LBP operators.

2.4. Pearson correlation coefficient

Pearson correlation coefficient (PCC) is a measurement of linear correlation between two random real-valued variables. It is the first and the most widely used measures of correlation. Pearson correlation coefficient is defined as the product of the covariance of two variables X and Y divided by their standard deviation which means the calculation process of Pearson correlation coefficient includes normalization, therefore, there is no requirement for the range of values between different variables, and the final obtained correlation measures the trend, while the difference in scale of different variables is removed in the calculation process. Its calculation formula [Eq. (1)] was proposed by Rodgers [37].

$$r_{xy} = \frac{\sum (x_i - \bar{x}) \sum (y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2} \sqrt{\sum (y_i - \bar{y})^2}} \quad (1)$$

The coefficient r_{xy} ranges from -1 to 1. When r_{xy} is greater than 0, x and y are positively correlated, and when r_{xy} is less than 0, x and y are negatively correlated. The greater the absolute value of the coefficient r_{xy} , the stronger the correlation. Normally, when the absolute value of r_{xy} is between 0.8 and 1.0, the correlation is extremely strong, between 0.6 and 0.8 represents a strong correlation, between 0.4 and 0.6 means a moderate correlation, and between 0.2 and 0.4, there is a weak correlation. Between 0.0 and 0.2, x and y are considered to be very weakly correlated or uncorrelated.

In this study, we used Pearson correlation coefficient to quantify the degree of correlation between the two sets of image texture features and polarization parameters.

2.5. Schematic of the study

Figure 3 summarizes the schematic of the study. In the preprocessing part, thirty intensity images and H&E images were first registered, and the regions of interest (900*1000 pixels) with the same size, same resolution and pixel-level registration were cut out. After obtaining the Mueller matrices using the 30 intensity images, polarization parameter images can be derived by the functions of Mueller matrix elements [8]. And then we calculated the texture features within 50*50 pixels window on the H&E image and slides the window on the entire image to obtain the texture parameters images. Perform threshold calculation on the texture parameter images with the maximum inter-class variance (Otsu algorithm) to segment the mask for obtaining targeted structures [38], then map the pixel position selected in the mask to both polarization parameter image and texture feature image, which ensures that the selected pixel locations contain microstructural information. A pair of Texture Feature-Polarization Parameter vectors is obtained in which pixel positions is corresponding, and the Pearson correlation coefficient between these two vectors is calculated to form a heat map. By analyzing the heat map, we could find the relationship between polarization parameters and texture feature parameters, which paves the way for the microstructural information of both modes to complement each other.

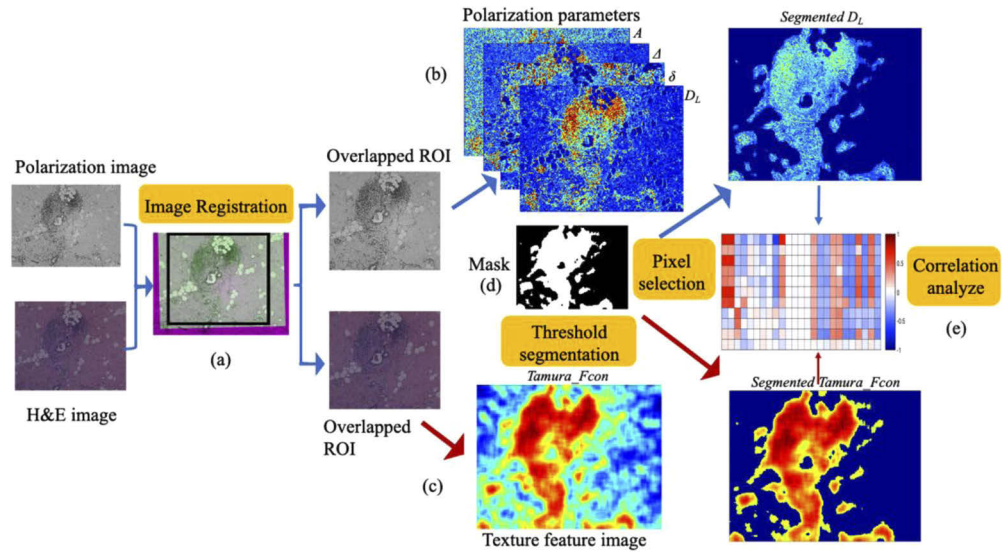


Fig. 3. Schematic of the study: (a) Preprocessing part that image registration between polarization image and H&E image. (b) Example of polarization parameters that derived from Mueller Matrices. (c) Example of texture feature image. (d) Mask formed by thresholding and segmentation from texture feature image. (e) Pearson correlation coefficient heatmap.

3. Results and discussion

We summarize the one-to-many associations of several pairwise texture feature-polarization parameters for different target microstructures in this section. For recognizing the cell nuclei structure, the Pearson correlation coefficient 0.6 was used as a threshold to filter out the highly correlation, image texture features $Tamura_Fcon$ and $Tamura_Frgh$ associated with polarization parameters of A , t_1 and D_L . For identifying the fiber structure, the Pearson correlation coefficient 0.4 was used as a threshold to filter out the highly correlated image texture features $Tamura_Flin$ connected with polarization parameters of δ and r_L . The performance of both texture features

and polarization parameters were tested on in total 15 breast ductal carcinoma tissue samples on multi-resolution cases.

3.1. Texture features of H&E images

In this study, the twenty-five texture feature images mentioned in section 2.3.2 were calculated. we selected 8 texture feature parameters among all the twenty-five texture feature parameters based on the parameter imaging contrast of cell nuclei or fiber structure in breast pathological tissues, i.e. the 8 texture features have potential for quantitative characterization of the target microstructures, as shown in Fig. 4.

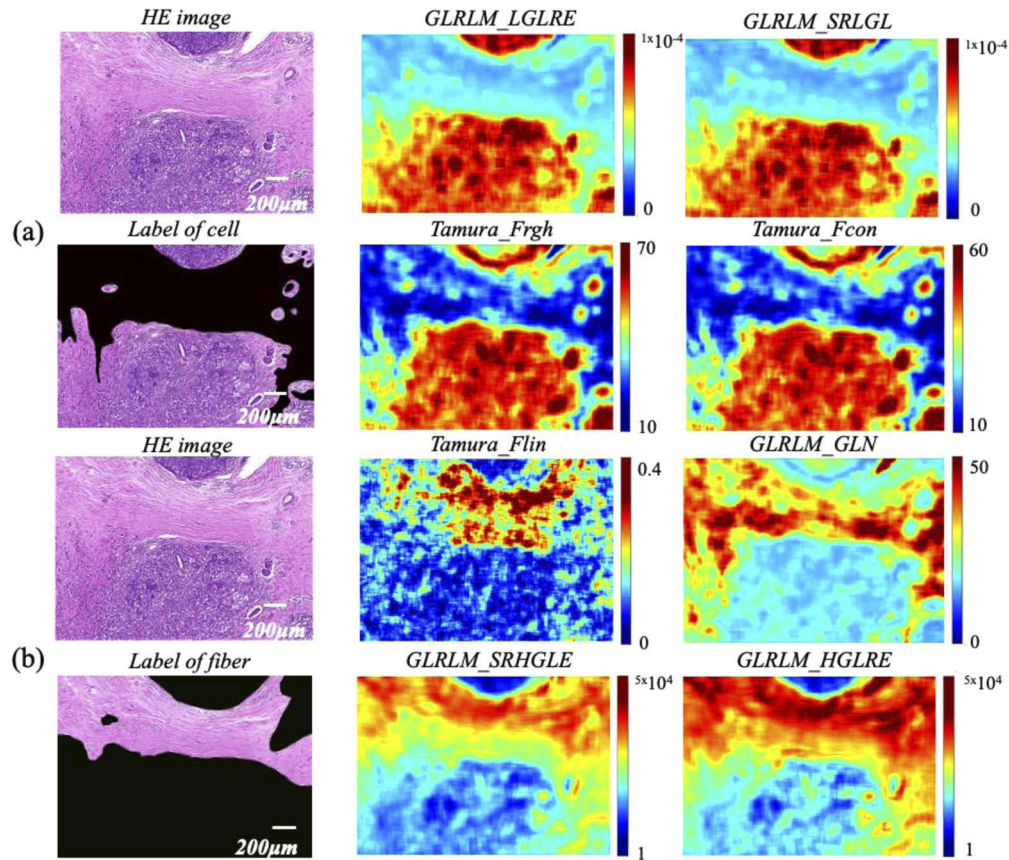


Fig. 4. Sets of texture feature images derived from H&E images of breast pathological tissue using image processing methods. Texture feature parameters-*GLRLM_LGLRE*, *GLRLM_SRLGRE*, *Tamura_Fcon* and *Tamura_Frgh*-are corresponding to the label of (a) cell nuclei, and texture feature parameters-*Tamura_Flin*, *GLRLM_GLN*, *GLRLM_SRHGRE*, and *GLRLM_HGLRE* -are corresponding to the label of (b) fiber tissue. The labels in this study come from experienced pathologists.

The image results for the texture features shown in Fig. 4 are calculated within a 50*50 pixel size window and assigned to the center point within the window, which is a process similar to convolution, and then sliding the window through the entire H&E image to obtain images of the texture features. Therefore, the value of each pixel in the texture feature image corresponds to the texture information in a small neighborhood around it in the original H&E image. The pixel values in the polarization parameter image represent the polarization feature in the corresponding

location of sample. Although the sliding window computation makes each pixel value in the texture feature image quantitatively characterize the texture information in that region, this approach pixelates information that would be extracted only within the region. Thus, the analysis of two datasets with one-to-one pixel-level correspondence becomes meaningful.

As shown in Fig. 4(a), the highlighted areas of *GLRLM_LGLRE*, *GLRLM_SRLGRE*, *Tamura_Fcon* and *Tamura_Frgh* all match the area marked by the label of the cell nuclei in breast ductal carcinoma tissue. The *Tamura_Fcon* feature reflects the sharpness of the edge and the period of repeating patterns, which has corresponding relationship with the concentrated area of cell nuclei under H&E-staining conditions. The *Tamura_Frgh* is a texture feature related to human visual tactile impression, which characterizes the degree of roughness or smoothness of the image. The cell nuclei area is visually uneven in the H&E-stained images, therefore, it is reasonable that the value of *Tamura_Frgh* is high in the cell nuclei regions of breast pathological tissues. The two features of *GLRLM_LGLRE* and *GLRLM_SRLGRE* are the statistical characteristics of low gray level, short run and low gray level in the gray run-length matrix respectively. We can observe from Fig. 4(a) that the high pixel values of the two statistical features distribute in the nuclei area, which has great potential for quantitative characterization of cell nuclei in breast pathological tissues.

As shown in Fig. 4(b), the highlighted areas of *Tamura_Flin*, *GLRLM_GLN*, *GLRLM_SRHGRE*, and *GLRLM_HGLRE* are corresponding with the area marked by the fiber label in the H&E stained images. The *Tamura_Flin* feature represents the degree of line-likeness of the shape of repeating patterns. The features of *GLRLM_SRHGRE* and *GLRLM_HGLRE* are the statistical characteristics of high gray level, short run and low gray level in the gray run-length matrix respectively, and *GLRLM_GLN* means non-uniformity of the image. Nuclei appear as dark dots in H&E-stained pathological sections. Thus, grayscale features of cell nuclei area can be summarized as low grayscale levels and short excursions, while texture features have high repeatability and clear boundaries. Different from the nuclei area, conversely, the gray-scale characteristic of the fibrous structure is high gray level, strong continuity and linear distribution. The characterization results indicate that the four texture features enable the identification of fiber structures in breast pathological slides.

3.2. Polarization parameters of tissues

The eleven polarization parameters PBPs mentioned in section 2.2 were calculated. PBPs images which have obvious imaging contrast on cell nuclei or fiber regions were selected, which may serve as powerful tools for quantitative characterization of target microstructures in breast ductal carcinoma tissues.

As shown in Fig. 5(a), The PBPs of D_L , t_1 , Δ , D and b have different characterization ability for the cell nuclei in breast histological samples. Parameters D and Δ are obtained by the MMPD method, D represents dichroism, and Δ represents the depolarization physical characteristics of sample. The parameter b calculated by the MMT method is related to the diagonal Mueller matrix elements m_{22} and m_{33} , which is closely related to the depolarization. t_1 represents the anisotropy, and A is calculated by t_1 and b , which represents normalized anisotropy. The black solid line drawn by pathologist in the H&E images of breast ductal carcinoma tissue labels the cell nuclei, which can be used as the ground truth for parameters characterization. By analyzing the comparison between the 2D images of PBPs and ground truth, we can conclude that the breast cell nuclei have anisotropy produced by dichroism, and depolarization related parameters images have good contrast between cell nuclei and the rest pathological structures. Therefore, the five selected PBPs were analyzed for the correlation with the image texture features which can be used for recognition of the breast cell nuclei.

Figure 5(b) shows that the PBPs of P_L , r_L , q_L , δ and θ have different ability to distinguish fiber tissue (marked by the red solid line) in breast pathological slides quantitatively. Parameters δ and

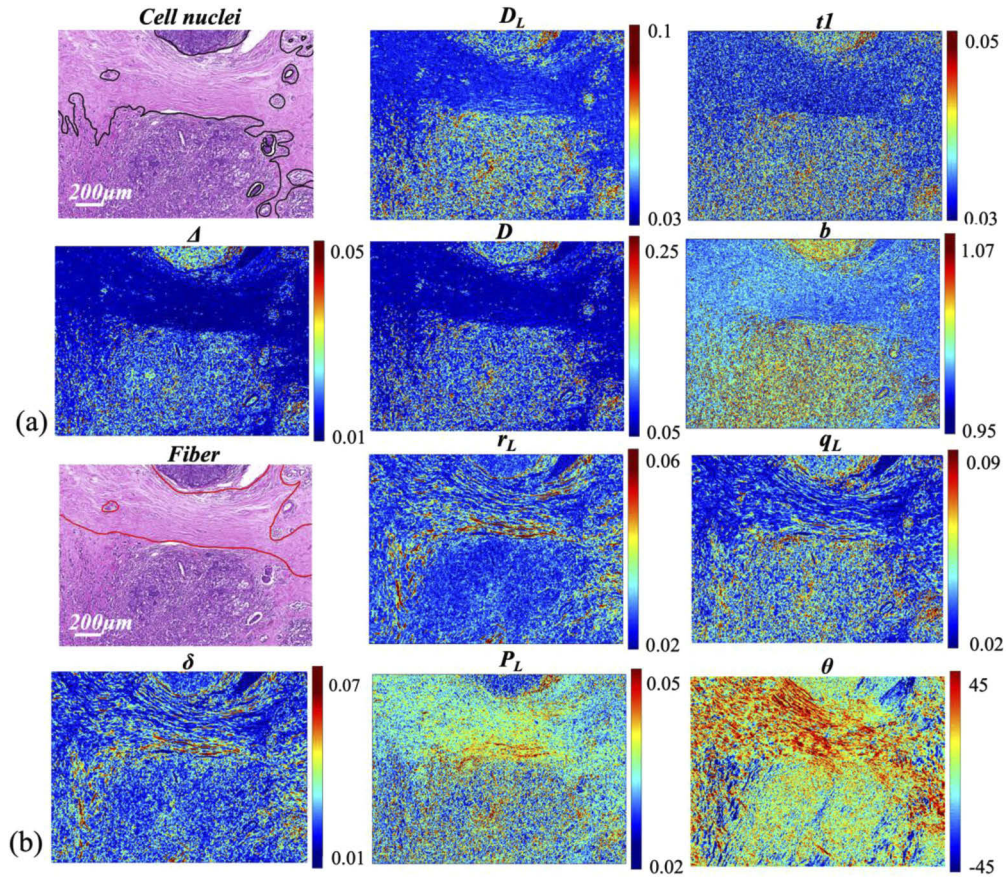


Fig. 5. Sets of polarization parameters images of breast pathological tissue. PBPs- D_L , t_1 , Δ , D and b - are correspond with the label of (a) cell nuclei (labelled by the black solid line in H&E stained image), and PBPs- P_L , r_L , q_L , δ and θ -are correspond with the label of (b) fiber tissue (marked by the red solid line in H&E stained image).

θ from MMPD represent linear retardation and angle of fast axis respectively. P_L , r_L and q_L are the modules of the left, right and lower edges of the Mueller matrix, which characterize column scattering, linear retardation and birefringence of the sample [14,19]. Therefore, the results of PBPs images indicate that the fibrous structure in breast pathological tissue have anisotropy produced by birefringence, resulting in the birefringence related parameters were selected to find the relationship with the texture features for the characterization of fiber structures in breast pathological tissue.

3.3. Building quantitative correspondence between polarization parameters and texture features

Figure 6 shows the Pearson coefficient (r_{xy}) correlation heatmap of pairwise association between the texture features and polarization parameters which were proved in above section that have ability to distinguish nuclei and fiber tissue. The horizontal axis is texture features, and the vertical axis is polarization parameters. Dark color means two sets of data is highly correlated. Viewed longitudinally from the perspective of polarimetry, PBPs with similar physical meanings are arranged in adjacent rows. We put the depolarization related parameters Δ and b , the overall anisotropy parameters A and t_1 , and the dichroism related parameters D and D_L in the top half

of the heatmap. These PBPs are sensitive to cell nuclei based on above analysis. Besides, the retardance related parameters (r_L , q_L , δ and θ) which have closely relationship with fiber structures are arranged in the bottom half of the heatmap. P_L can be considered as a dividing line for PBPs characterization between the nuclei and the fiber tissue. We can observe from heatmap that there are obvious positive and negative changes as well as different absolute values of coefficients above and below parameter P_L , indicating that PBPs sets characterizing the different target microstructures have significantly different correlation with the texture feature parameters. The texture parameters ($GLRLM_LGLRE$, $GLRLM_SRLGRE$, $Tamura_Fcon$ and $Tamura_Frgh$) which can be used for cell nuclei recognition based on above analysis have large correlation values with PBPs above P_L . The texture parameters $Tamura_Flin$, $GLRLM_GLN$, $GLRLM_SRHGRE$, and $GLRLM_HGLRE$ make great contributions for the characterization of fiber tissue in breast pathological slides, which have positive correlation coefficients with PBPs below parameter P_L , whereas are negative correlated with that above P_L . Of note, it can be found that polarization parameters with similar physical meanings will have relatively large differences when calculating correlation coefficients for the same texture features. This may be due to the fact that the information contained in the parameters decoded by the different methods are not exactly consistent. It is shown in Fig. 6 that r_{xy} between orientation of fast axis θ and sets of image texture features are very small, however, we can observe from Fig. 5 that the 2D image parameter θ has obvious imaging contrast on the fiber structures. Polarization parameter θ is related explicitly to the orientation angle of the sample. It is the statistical distribution features of θ , such as standard deviation, rather than the numerical value that has potential to quantitatively characterize the fiber structures [39]. Therefore, when the Pearson correlation coefficients are calculated using the value of angle parameter, θ does not correlate with the individual texture features.

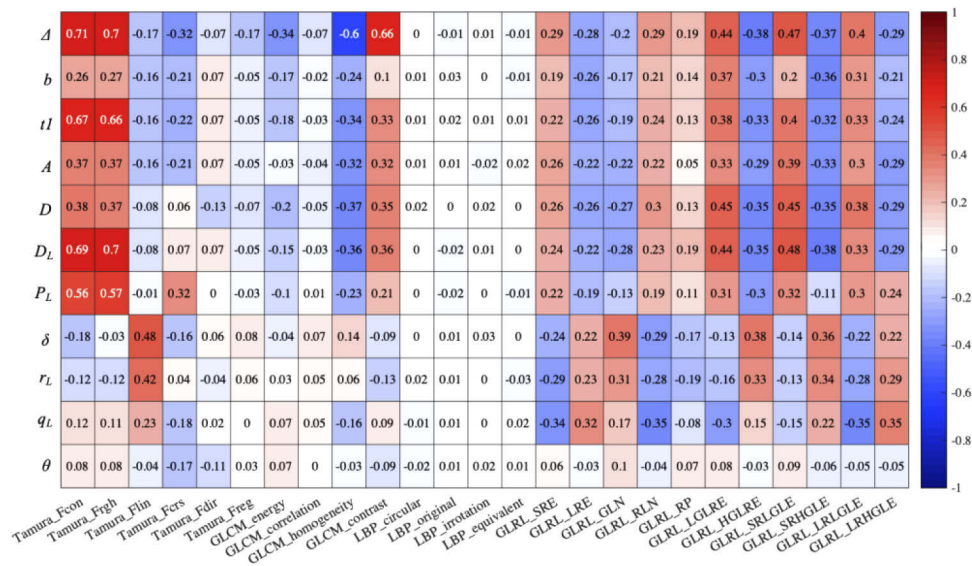


Fig. 6. Correlation heatmap. The correlation heat map shows the Pearson coefficient between the polarization parameters (y-axis) and the image texture features (x-axis). Red means positive correlation, blue means negative correlation, and white means uncorrelation or un-statistically significant correlation.

We can observe from Fig. 6 that the r_{xy} of pairwise association between the texture features and polarization parameters used for characterization of cell nuclei structures was higher than that of the fiber regions. Thus, for screening the image texture feature-polarization parameter with high

correlation, we took the baseline of the r_{xy} for different target structures into consideration. The baseline is calculated under an extreme condition in which the resolutions of H&E-stained images and the polarization parameter images are reduced significantly using a mean filter sliding window with 100*100 size, then texture feature images of the blurred H&E images were calculated within a 50*50 sliding window. The correlation of the image texture feature-polarization parameter under this extreme condition tends to be stable and minimal, with the baseline of r_{xy} for cell nuclei is about 0.2 and the baseline of r_{xy} for fiber structure is near to 0. In Section 2.4, we have introduced the smallest r_{xy} of moderate correlation is 0.4. Considering the baseline of r_{xy} for different target structures under an extreme condition and the smallest r_{xy} of moderate correlation, the threshold is set to the sum of the two. Therefore, for nuclei structures, r_{xy} of 0.6 was used as a threshold to screen out highly correlated image texture features *Tamura_Fcon* and *Tamura_Frgh* with depolarization parameter Δ , polarization parameters t_1 and D_L , and for fiber structures, r_{xy} of 0.4 was used as a threshold to screen out highly correlated image texture features *Tamura_Flin* with polarization parameters δ and r_L .

Table 3 summarizes all the eight texture features which were highly and significantly associated with polarization parameters with high Pearson coefficient. Table 4 summarizes the texture features that are significantly related to the polarization parameters (with r_{xy} of cell nuclei ≥ 0.6 , r_{xy} of fiber ≥ 0.4) in the heat map from the perspective of the polarization parameters. This summary provides potential physical characteristics for the image parameters when characterizing the same microstructures in pathological samples, and selection methods for polarization parameters by image feature parameters.

Table 3. Summary of representative texture features

| Tissue | Texture features | r_{xy} | Feature explanation |
|-------------|---------------------|----------|--|
| Cell nuclei | <i>Tamura_Fcon</i> | 0.71 | Quantitatively describe sharpness of edges, period of cell repeating patterns. |
| | <i>Tamura_Frgh</i> | 0.7 | Quantitatively describe total energy of changes in specific area. |
| | <i>GLRLM_LGRLE</i> | 0.45 | Distribution of the low grey-level runs. |
| | <i>GLRLM_SRLGLE</i> | 0.38 | Distribution of the short homogeneous runs with low grey-levels. |
| Fiber | <i>Tamura_Flin</i> | 0.48 | Quantitatively describe the shape of texture element. |
| | <i>GLRLM_GLN</i> | 0.39 | Describe non-uniformity of the grey-levels. |
| | <i>GLRLM_HGRLE</i> | 0.38 | Distribution of the high grey-level runs. |
| | <i>GLRLM_SRHGLE</i> | 0.36 | Distribution of the short homogeneous runs with high grey-levels. |

Table 4. Summary of representative polarization parameters

| Tissue | Polarization Parameters | Numbers and details related to texture features | Physics explanation of parameters |
|-------------|-------------------------|---|-----------------------------------|
| Cell nuclei | Δ | 4 (<i>Tamura_Fcon</i> , <i>Tamura_Frgh</i> , <i>GLCM_Hmg</i> , <i>GLCM_Con</i>) | Depolarization. |
| | D_L | 2 (<i>Tamura_Fcon</i> , <i>Tamura_Frgh</i>) | Linear dichroism. |
| | t_1 | 2 (<i>Tamura_Fcon</i> , <i>Tamura_Frgh</i>) | Anisotropy degree. |
| Fiber | δ | 1 (<i>Tamura_Flin</i>) | Linear retardation. |
| | r_L | 1 (<i>Tamura_Flin</i>) | Linear retardation. |

These highly associated relations may lead to potential conjectures about image texture features with specific microstructures or even physical quantity. Texture features *Tamura_Fcon* and *Tamura_Frgh* can quantitatively describe cell nuclei in H&E pathological images of breast tissues, and PBPs t_1 , D_L and depolarization Δ enable the quantitative characterization of the same target

microstructure. It may imply that on H&E stained cases, texture features *Tamura_Fcon* and *Tamura_Frgh* have close relationship with physical meanings of depolarization, anisotropy and dichroism. The aligned fiber tissue is highly line-like that can be characterized by texture feature *Tamura_Flin*. Meanwhile, polarization parameters δ and r_L are sensitive to fibrous structure. Thus, for H&E-stained cases, texture feature *Tamura_Flin* may have physical feature of linear retardation. On the other hand, after finding the relationship between PBPs and texture features, we could predict sample's specific polarization features without Mueller matrix measurement only by calculating the corresponding image texture parameters.

3.4. Validation

3.4.1. Validation of the stability of texture features

By applying different sizes of average filter before input the H&E image, we verified the stability of texture features on multi-resolution cases. During the validation process, 15 samples were tested. Specifically, we compared the results obtained from thresholding texture feature images at different resolution cases with the gold standard, and the accuracy was calculated as Eq. (2). The ability of this texture feature to recognize the target structure is thus quantified. The accuracy of image texture features to identify target structures significantly decreases with decreasing resolution. As shown in Fig. 7, for the characterization of cell nuclei, the accuracy of texture feature *Tamura_Fcon* decreased from 0.87 (without average filtering) to 0.73 (with 20*20 average filtering) gradually, and that of texture feature *Tamura_Frgh* decreased from 0.89 to 0.76. For the recognition of fibrous tissues, texture feature *Tamura_Flin* decreased from 0.83 to 0.69 as the image resolution decreases. These results suggest that image texture features have good and stable identification of target structures at high resolution, however, this identification is heavily resolution dependent.

3.4.2. Validation of the stability of polarization parameters

Likewise, by applying different sizes of average filter directly on polarization parameter images, we verified the stability of polarization parameters on multi-resolution cases. For the characterization of cell nuclei tissue, the accuracy of the depolarization parameter Δ was 0.87 in the no-means-filtered case, while it remained at 0.85 in the 20*20-means-filtered case, decreasing only by 0.02. Meanwhile, the accuracy of the polarization parameter D_L was kept around 0.87 with only slight fluctuations, and that of the polarization parameter t_1 slightly decreased from 0.86 to 0.82. For the identification of fibrous tissues, the accuracy of the polarization parameter r_L was maintained between 0.85 and 0.86, and the accuracy of the polarization parameter δ decreased by 0.02 after 0.83 as the resolution decreasing. These results show that the polarization parameters have good and stable identification of the target structure at multi-resolution cases.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \% \quad (2)$$

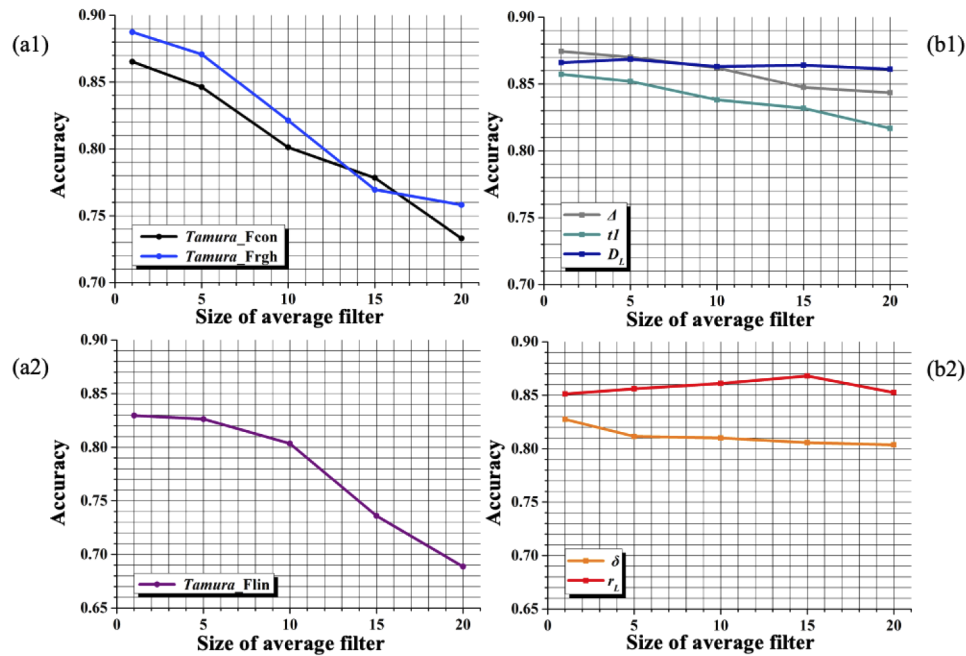


Fig. 7. Accuracy on the characterization of the target microstructures by Texture features-Polarization parameters at multi-resolution cases. (a1) Texture features *Tamura_Fcon* and *Tamura_Frgh* for the identification of cell nuclei. (a2) Texture feature *Tamura_Flin* for the recognition of fiber tissue. (b1) Polarization parameters tI , D_L and depolarization Δ for the characterization of cell nuclei. (b2) Polarization parameters δ and r_L for the characterization of fiber tissue.

4. Conclusion

In this study, we proposed an approach that can successfully validate the hypothesis associations between image texture features and polarization parameters, which reflect the tissue microstructure classification on breast ductal carcinoma. After extensively calculating Pearson correlation coefficients for 25 image texture features and 11 polarization parameters, we screened stable and highly correlated texture feature-polarization parameter correspondences for different target microstructures.

Polarization parameters tI , D_L and depolarization parameter Δ can be used to quantitatively characterize cell nuclei tissue, which related to image texture features *Tamura_Fcon* and *Tamura_Frgh*. Polarization parameters δ and r_L association with image texture feature *Tamura_Flin*, which have ability to recognize fibrous tissue in breast ductal carcinoma samples. We used 15 breast ductal carcinoma samples for validation at multiple resolutions and found that both texture features and polarization parameters performed satisfactorily in identifying target structures at high resolution, but at low resolution, the accuracy of texture features decreased significantly with decreasing resolution, while polarization parameters showed stable accuracy. This study realizes the potential for image texture features to provide complementary information to polarization parameters at high resolution, and for polarization imaging methods to provide specific physical meanings to light intensity images on histopathological samples.

Funding. Shenzhen Municipal Science and Technology Innovation Council (JCYJ20160818143050110, JCYJ20170412170814624); National Natural Science Foundation of China (11974206, 61527826).

Acknowledgments. The authors are also thankful to Yuanhuan Zhu from Tsinghua-Berkeley Shenzhen Institute for useful advices, and Tongyu Huang and Jia Dong from the Tsinghua Shenzhen International Graduate School for the assistants.

Disclosures. The authors declare no conflicts of interest.

References

1. M. Veta, J. P. W. Pluim, P. J. van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: a review," *IEEE Trans. Biomed. Eng.* **61**(5), 1400–1411 (2014).
2. M. W. Conklin, J. C. Eickhoff, K. M. Riching, C. A. Pehlke, K. W. Eliceiri, P. P. Provenzano, A. Friedl, and P. J. Keely, "Aligned collagen is a prognostic signature for survival in human breast carcinoma," *Am. J. Pathol.* **178**(3), 1221–1232 (2011).
3. C. Lu, M. Ji, Z. Ma, and M. Mandal, "Automated image analysis of nuclear atypia in high-power field histopathological image," *J. Microsc.* **258**(3), 233–240 (2015).
4. B. Jones, G. Thomas, J. Sprenger, S. Nofech-Mozes, M. Khorasani, and A. Vitkin, "Peri-tumoural stroma collagen organization of invasive ductal carcinoma assessed by polarized light microscopy differs between OncotypeDX risk group," *J. Biophotonics* **13**(11), e202000188 (2020).
5. G. Chen, G. Liu, X. Zhu, M. Liu, E. Zhang, J. Li, K. Zhang, and L. Lin, "A novel method for describing texture of scar collagen using second harmonic generation images," *IEEE Photonics J.* **9**(2), 1–13 (2017).
6. M. Borovkova, L. Trifonyuk, V. Ushenko, O. Dubolazov, and I. Meglinski, "Mueller-matrix-based polarization imaging and quantitative assessment of optically anisotropic polycrystalline networks," *PLoS One* **14**(5), e0214494 (2019).
7. A. Ushenko, Y. Sarkisova, A. Dubolazov, Y. Ushenko, Y. Tomka, A. Karachevtsev, M. Sidor, A. Prydiy, V. Dvorjak, and K. Tymchuk, "Informativeness of the medical-physical method for analyzing the distributions of the degree of local depolarization of microscopic images of histological sections of the intestinal wall for the diagnosis of dolechosisigma," *Proc. SPIE* **11369**, 113691O (2020).
8. S. Y. Lu and R. A. Chipman, "Interpretation of Mueller matrices based on polar decomposition," *J. Opt. Soc. Am. A* **13**(5), 1106–1113 (1996).
9. H. He, R. Liao, N. Zeng, P. Li, Z. Chen, X. Liu, and H. Ma, "Mueller matrix polarimetry—an emerging new tool for characterizing the microstructural feature of complex biological specimen," *J. Lightwave Technol.* **37**(11), 2534–2548 (2019).
10. P. Li, H. R. Lee, S. Chandel, C. Lotz, F. Kai Groeber-Becker, S. Dembski, R. Ossikovski, H. Ma, and T. Novikova, "Analysis of tissue microstructure with Mueller microscopy: logarithmic decomposition and Monte Carlo modeling," *J. Biomed. Opt.* **25**(1), 015002 (2020).
11. H. R. Lee, T. S. H. Yoo, P. Li, C. Lotz, F. K. Groeber-Becker, S. Dembski, E. Garcia-Caurel, R. Ossikovski, and T. Novikova, "Mueller microscopy of anisotropic scattering media: theory and experiments," *Proc. SPIE* **10677**, 1067718 (2018).
12. E. Du, H. He, N. Zeng, M. Sun, Y. Guo, J. Wu, S. Liu, and H. Ma, "Mueller matrix polarimetry for differentiating characteristic features of cancerous tissues," *J. Biomed. Opt.* **19**(7), 076013 (2014).
13. Y. Dong, J. Wan, L. Si, Y. Meng, Y. Dong, S. Liu, H. He, and H. Ma, "Deriving polarimetry feature parameters to characterize microstructural features in histological sections of breast tissues," *IEEE Transactions on Biomedical Engineering*, 10.1109/TBME.2020.3019755 (2020).
14. Y. Dong, J. Qi, H. He, C. He, S. Liu, J. Wu, D. S. Elson, and H. Ma, "Quantitatively characterizing the microstructural features of breast ductal carcinoma tissues in different progression stages by Mueller matrix microscope," *Biomed. Opt. Express* **8**(8), 3643 (2017).
15. V. Dremine, D. Anin, O. Sieryi, M. Borovkova, J. Näpänkangas, I. Meglinski, and A. Bykov, "Imaging of early stage breast cancer with circularly polarized light," *Proc. SPIE* **11363**, 1136304 (2020).
16. C. He, H. He, J. Chang, Y. Dong, S. Liu, N. Zeng, Y. He, and H. Ma, "Characterizing microstructures of cancerous tissues using multispectral transformed Mueller matrix polarization parameters," *Biomed. Opt. Express* **6**(8), 2934–2945 (2015).
17. Y. Dong, H. He, C. He, J. Zhou, N. Zeng, and H. Ma, "Characterizing the effects of washing by different detergents on the wavelength-scale microstructures of silk samples using Mueller matrix polarimetry," *Int. J. Mol. Sci.* **17**(8), 1301–1314 (2016).
18. H. He, N. Zeng, E. Du, Y. Guo, D. Li, R. Liao, and H. Ma, "A possible quantitative Mueller matrix transformation technique for anisotropic scattering media/Eine mögliche quantitative Müller-Matrix-Transformations-Technik für anisotrope streuende Medien," *Photonics Lasers Med.* **2**, 129–137 (2013).
19. P. Li, D. Lv, H. He, and H. Ma, "Separating azimuthal orientation dependence in polarization measurements of anisotropic media," *Opt. Express* **26**(4), 3791–3800 (2018).
20. G. Lu, D. Wang, X. Qin, S. Muller, and B. Fei, "Histopathology feature mining and association with hyperspectral imaging for the detection of squamous neoplasia," *Sci. Rep.* **9**(1), 17863 (2019).
21. Y. Wang, H. He, J. Chang, N. Zeng, S. Liu, M. Li, and H. Ma, "Differentiating characteristic microstructural features of cancerous tissues using Mueller matrix microscope," *Micron* **79**, 8–15 (2015).
22. D. H. Goldstein, "Mueller matrix dual-rotating retarder polarimeter," *Appl. Opt.* **31**(31), 6676–6683 (1992).

23. D. H. Goldstein and R. A. Chipman, "Error analysis of a Mueller matrix polarimeter," *J. Opt. Soc. Am.* **7**(4), 693–700 (1990).
24. T. Liu, M. Lu, B. Chen, Q. Zhong, J. Li, H. He, H. Mao, and H. Ma, "Distinguishing structural features between Crohn's disease and gastrointestinal luminal tuberculosis using Mueller matrix derived parameters," *J. Biophotonics* **12**(12), e201900151 (2019).
25. P. Li, A. Tariq, H. He, and H. Ma, "Characteristic Mueller matrices for direct assessment of the breaking of symmetries," *Opt. Lett.* **45**(3), 706–709 (2020).
26. H. J. W. L. Aerts, E. R. Velazquez, R. T. H. Leijenaar, C. Parmar, and P. Lambin, "Corrigendum: Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nat. Commun.* **5**(1), 4406 (2014).
27. A. D. Belsare, M. M. Mushrif, M. A. Pangarkar, and N. Meshram, "Classification of breast cancer histopathology images using texture feature analysis," *TENCON 2015 - 2015 IEEE Region 10 Conference* (2016).
28. M. Jenkinson and S. Smith, "A global optimisation method for robust affine registration of brain images," *Med. Image Anal.* **5**(2), 143–156 (2001).
29. H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.* **8**(6), 460–473 (1978).
30. R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.* **SMC-3**(6), 610–621 (1973).
31. T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," *Computer Vision - ECCV 2004. Proceedings Part I* (2004).
32. T. Ojala, M. Pietikainen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," *Conference A: Computer Vision & Image Processing. Proceedings of the 12th IAPR International Conference on IEEE* (1994).
33. T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognition* **29**(1), 51–59 (1996).
34. M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graph. Image Process.* **4**(2), 172–179 (1975).
35. A. Chu, C. M. Sehgal, and J. F. Greenleaf, "Use of gray value distribution of run lengths for texture analysis," *Pattern Recognition Lett.* **11**(6), 415–419 (1990).
36. B. V. Dasarathy and E. B. Holder, "Image characterizations based on joint gray level-run length distributions," *Pattern Recognition Lett.* **12**(8), 497–502 (1991).
37. J. Lee Rodgers and W. A. Nicewander, "Thirteen ways to look at the correlation coefficient," *Am. Stat.* **42**(1), 59–66 (1988).
38. N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.* **9**(1), 62–66 (1979).
39. Y. Dong, H. He, W. Sheng, J. Wu, and H. Ma, "A quantitative and non-contact technique to characterise microstructural variations of skin tissues during photo-damaging process based on Mueller matrix polarimetry," *Sci. Rep.* **7**(1), 14702 (2017).